1-1-2007

Discrete Approximations, Relaxation, and Optimization of One-Sided Lipschitzian Differential Inclusions in Hilbert Spaces

Tzanko Donchev
University of Architecture and Civil Engineering, Sofia, Bulgaria, tzankodd@gmail.com

Elza Farkhi
Tel Aviv University, Israel, elza@post.tau.ac.il

Boris S. Mordukhovich
Wayne State University, boris@math.wayne.edu

DISCRETE APPROXIMATIONS, RELAXATION, AND OPTIMIZATION OF ONE-SIDED LIPSCHITZIAN DIFFERENTIAL INCLUSIONS IN HILBERT SPACES

TZANKO DONCHEV, ELZA FARKHI, and
BORIS. S. MORDUKHOVICH

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2007 Series
#1

This research was partly supported by the National Science Foundation and the Australian Research Council

DISCRETE APPROXIMATIONS, RELAXATION, AND OPTIMIZATION OF ONE-SIDED LIPSCHITZIAN DIFFERENTIAL INCLUSIONS IN HILBERT SPACES

TZANKO DONCHEV

Department of Mathematics, University of Architecture and Civil Engineering
1046 Sofia, Bulgaria; tzankodd@gmail.com

ELZA FARKHI

School of Mathematical Sciences, Tel Aviv University
69978 Tel Aviv, Israel; elza@post.tau.ac.il

and

BORIS S. MORDUKHOVICH [1]

Department of Mathematics, Wayne State University
Detroit, Michigan 48202, USA; boris@math.wayne.edu

Dedicated to Arrigo Cellina in honor of his 65th birthday

Abstract. We study discrete approximations of nonconvex differential inclusions in Hilbert spaces and dynamic optimization/optimal control problems involving such differential inclusions and their discrete approximations. The underlying feature of the problems under consideration is a modified *one-sided Lipschitz* condition imposed on the right-hand side (i.e., on the velocity sets) of the differential inclusion, which is a significant improvement of the conventional Lipschitz continuity. Our main attention is paid to establishing efficient conditions that ensure the *strong approximation* (in the $W^{1,p}$-norm as $p \geq 1$) of feasible trajectories for the one-sided Lipschitzian differential inclusions under consideration by those for their discrete approximations and also the *strong convergence* of optimal solutions to the corresponding dynamic optimization problems under discrete approximations. To proceed with the latter issue, we derive a new extension of the Bogolyubov-type relaxation/density theorem to the case of differential inclusions satisfying the modified one-sided Lipschitzian condition. All the results obtained are new not only in the infinite-dimensional Hilbert space framework but also in finite-dimensional spaces.

Keywords. Differential Inclusions; Discrete approximations; One-sided Lipschitz condition; Optimal control; Relaxation stability; Strong convergence of optimal solutions

Mathematics Subject Classifications (2000). 49J24, 49M25, 90C99

Corresponding author: Boris Mordukhovich, Phones: (313)577-3193 (of), (734)995-2659 (h); Fax: (313)577-7596; Email: boris@math.wayne.edu

[1]Research of this author was partially supported by the USA National Science Foundation under grants DMS-0304989 and DMS-0603846 and by the Australian Research Council under grant DP-0451168.

1 Introduction

This paper is devoted to the study of *differential inclusions* given in the form

$$\dot{x}(t) \in F\big(t, x(t)\big) \text{ for a.e. } t \in T := [0,1], \quad x(0) = x_0 \in H, \tag{1.1}$$

where H is a *Hilbert space*, and where $F\colon T \times H \rightrightarrows H$ is a set-valued mapping with nonempty *compact* values (some results hold also with no compactness assumption; see Remark 4.4 for more discussions). It is well known that the differential inclusion description (1.1) is important for its own sake and covers many other conventional and nonconventional models involving dynamical systems in finite and infinite dimensions; see, e.g., [1, 2, 5, 12, 15, 17] and the references therein. In particular, differential inclusions (1.1) extend control systems

$$\dot{x}(t) = f(t, x, u), \quad u \in U(t, x), \tag{1.2}$$

where the control region $U(t, x)$ can depend on the state variable x, which is a challenging issue in control theory and applications.

The primary purpose of this paper is to study *discrete approximations* of differential inclusions and certain *dynamic optimization* problems associated with them. These topics have been addressed in many publications, mostly in finite-dimensional spaces; see, e.g., surveys [12, 15] and the recent book [17] with more references and discussions. The vast majority of publications in these directions impose the classical *Lipschitz continuity* of the mapping F in x, which seems to be restrictive for a number of applications.

In this paper we systematically replace the Lipschitz continuity by a certain *modified one-sided Lipschitzian (MOSL) property* of F in x, which is an essentially weaker assumption; see more discussions below. Differential inclusions and their discrete approximations under the more conventional one-sided Lipschitz (OSL) condition have been already studied by the first two authors in papers [7, 8, 9, 10, 11] mostly devoted to qualitative theory of OSL differential inclusions and the possibility to *uniformly* approximate solutions sets to OSL inclusions (1.1) by corresponding solution sets to their discretized counterparts.

The scope and results of this paper are fully different from the previous developments. Our main efforts are to establish the *strong approximation* (in the $W^{1,p}$-norm as $p \geq 1$) of feasible trajectories for MOSL differential inclusions (1.2) by those for their discrete approximations and also to justify the *strong* $W^{1,p}$*-convergence* of optimal solutions to the associated problems of *dynamic optimization/optimal control* under discrete approximations. The results obtained in this paper extend, to the case of MOSL differential inclusions in finite-dimensional and Hilbert spaces, the corresponding developments of the third author [16, 17, 18] for differential inclusions satisfying the classical Lipschitz condition.

Another achievement of this paper, motivated by applications to the convergence of discrete approximations in optimal control while certainly significant for its own sake, is establishing a Bogolyubov-type *relaxation/density* theorem for differential inclusions satisfying the MOSL condition. The latter result is known to *hold* for Lipschitzian differential inclusions and to *fail* for OSL ones. All the results obtained in this paper seem to be *new* in both *finite-dimensional* and *infinite-dimensional* settings.

The rest of the paper is organized as follows. In Section 2 we formulate and discuss the standing assumptions and then present some preliminary material, which is broadly used for deriving the main results of the paper.

Section 3 is devoted to the study of relationships between *solution sets* to MOSL differential inclusions and those to their *discrete* approximations constructed via the *Euler finite-difference scheme* as well as to related *semi-discrete* approximations of (1.1). The main results justify, under the MOSL property of $F(t,\cdot)$, the possibility of the *strong* $W^{1,p}$*-norm approximation* of any feasible trajectory for (1.1) by those for its discrete and semi-discrete counterparts constructed in what follows.

In Section 4 we derive certain *density/relaxation stability* results of the Bogolyubov type concerning relationships between trajectories to the original MOSL differential inclusion coupled with an integral cost functional and the corresponding relaxed/convexified counterpart. The results obtained seem to be new in the extensive literature on relaxation stability and related topics (e.g., Young measures) for variational problems; they are sensitive even to a slight change of assumptions. Applying the technique developed in the proof of the main density theorem, we justify in this section a new (different from that in Section 3) version of the strong convergence theorem for discrete approximations imposing milder time-dependence assumptions on the initial data.

The concluding Section 5 deals with *discrete approximations* of *dynamic optimization* Bolza-type problems for nonconvex MOSL differential inclusions. It contains a major result of the paper justifying the *strong* $W^{1,p}$*-convergence* of optimal solutions for the discrete approximation problems to the given optimal solution (actually an arbitrary *local minimizer* of the "relaxed intermediate" and strong types) for the continuous-time generalized Bolza problem under consideration. We also establish general conditions (both *necessary* and *sufficient*) for the *value convergence* of discrete approximations of the generalized Bolza problem for MOSL differential inclusions. The results obtained in this section significantly improve known results in this direction by weakening assumptions on the initial data dependence with respect to both the state and time variables. The proofs given in this section are essentially based on the previous results of the paper on strong approximation and relaxation stability for MOSL differential inclusions.

Our notation is basically standard, with some special symbols explained in the text where they are introduced. Note that $I\!B$ stands for the closed unit ball of the space in question and that, given a subset Ω of the Hilbert space H under consideration with its norm denoted by $|\cdot|$, the symbols $\overline{\Omega}$ and $\mathrm{co}\,\Omega$ signify the closure of Ω and the convex hull of Ω, respectively; $I\!N := \{1, 2, \ldots\}$ stands for all the collection of natural numbers. Let us also mention that the constant $C > 0$ used in the proofs and various estimates throughout the paper is commonly a *generic constant*.

2 Basic Assumptions and Preliminaries

In this section we impose and discuss the underlying assumptions on the set-valued mapping F from (1.1) standing throughout the whole paper and then present several known facts on differential inclusions formulated in two lemmas, which are essential for proving the main results of the paper.

Given two closed and bounded sets $\Omega_1, \Omega_2 \subset Z$ in some Banach space Z with the norm $\|\cdot\|$, recall that the *Hausdorff distance* $d_Z(\Omega_1, \Omega_2)$ between them in Z is defined by

$$d_Z(\Omega_1, \Omega_2) := \max\big\{ \sup_{z\in\Omega_1} \operatorname{dist}(z;\Omega_2),\ \sup_{y\in\Omega_2} \operatorname{dist}(y;\Omega_1)\big\} \ \text{ with } \ \operatorname{dist}(z;\Omega) := \inf_{\omega\in\Omega} \|z-\omega\|.$$

As usual, a set-valued mapping $G\colon Y \rightrightarrows Z$ between two Banach spaces is *continuous* on some set $\Omega \subset Y$ if it is continuous on Ω with respect to the Hausdorff distance; it is *Lipschitz continuous* on Ω with modulus $L \geq 0$ if

$$d_Z\big(G(y_1), G(y_2)\big) \leq L\|y_1 - y_2\| \ \text{ whenever } \ y_1, y_2 \in \Omega. \tag{2.1}$$

Recall further that a nonautonomous mapping $G\colon T \times Y \rightrightarrows Y$ is *almost continuous* on $T \times \Omega$ if for every $\varepsilon > 0$ there is a compact set $T_\varepsilon \subset T$ such that $\operatorname{mes}(T \setminus T_\varepsilon) < \varepsilon$ and $G(\cdot,\cdot)$ is continuous on $T_\varepsilon \times \Omega$. We refer the reader to the book [5] for the standard definitions of *lower semicontinuity* (LSC) and *upper semicontinuity* (USC) and their similarly defined *almost LSC* and *almost USC* counterpart. Furthermore, in [5] the reader can find the conventional definitions of *measurable* and *strongly measurable* multifunctions; note that the latter notions agree when the range space is separable.

Now we formulate the following *standing assumptions* imposed on the set-valued mapping $F\colon T \times H \rightrightarrows H$ in our differential inclusion (1.1) defined on the Hilbert space H considering, unless otherwise stated, only mappings with nonempty and *compact* values.

(A1) $F\colon T \times H \rightrightarrows H$ is almost continuous and bounded on bounded sets.

(A2) There exist a constant $L \in I\!R$ and an almost continuous function $f\colon T \times I\!R_+ \to I\!R_+$ with the following properties:
(i) $f(t, 0) \equiv 0$, and it is bounded on bounded sets;
(ii) Given any $x_1, x_2 \in H$ and $y_1 \in F(t, x_1)$, there exists $y_2 \in F(t, x_2)$ such that

$$\langle x_1 - x_2, y_1 - y_2\rangle \leq L|x_1 - x_2|^2 \ \text{ and } \ |y_1 - y_2| \leq f\big(t, |x_1 - x_2|\big) \tag{2.2}$$

for almost all $t \in T$.

Note that the property described by the first inequality in (2.2) is known, for time-independent mappings, as the *one-sided Lipschitz* (*OSL*) *property* of $F(t,\cdot)$; see the references in Section 1 with more discussions given therein. The full property (A2) is a strengthened version of assumption ($H4$) from [7]; we call this new version the *modified one-sided Lipschitz* (*MOSL*) *property* of multifunctions. It obviously holds when $F(t,\cdot)$ satisfies the classical Lipschitz condition (2.1), while the *measurable* time dependence of $F(\cdot, x)$ is covered by (A2) due to Lusin/Scorza-Dragoni's type theorems for set-valued and single-valued mappings; see, e.g., [5, 20]. Observe that, in contrast to (2.1), the constant L in (2.2) is *not required to be positive.* This allows us to significantly extend the class of MOSL mappings in comparison with Lipschitz continuous mappings conventionally considered in the theory and applications of discrete approximations and optimization for differential inclusions. A simple example of a non-Lipschitzian (in the classical sense) function satisfying

(A2) is $-x^{1/3}$. A more involved situation when the MOSL property holds while $F(t,\cdot)$ is not Lipschitz continuous is given by the two-dimensional differential inclusion

$$\begin{cases} \dot{x}_1 \in \{-1,1\}, \ x_1(0)=0, \\ \dot{x}_2 = |x_1| - \operatorname{sign}(x_2)\sqrt{|x_2|}, \ x_2(0)=0. \end{cases} \tag{2.3}$$

On the other hand, it is easy to check that the MOSL property implies the uniform continuity of $F(t,\cdot)$. It is definitely stronger (more restrictive) than the standard OSL property used in the literature. This stronger assumption, together with (A1), allows us to establish here essentially stronger results than those known for OSL differential inclusions, with *no* imposing the full Lipschitz continuity (2.1). In particular, we justify the *strong* $W^{1,2}$*-norm approximation* of solutions to (1.1) by discrete and semi-discrete trajectories in Section 3 as well as the Bogolyubov-type *relaxation/density* results of Section 4. The latter result is known to *fail* under the standard OSL property; see, e.g., [4, Example 1.3].

In what follows, along with the original differential inclusion (1.1) we consider its *relaxation*, which is obtained from (1.1) by using the *convex closure* of $F(t,x)$:

$$\dot{x}(t) \in \overline{\text{co}}\, F\big(t,x(t)\big) \ \text{ for a.e. } \ t \in T, \quad x(0)=x_0 \in H. \tag{2.4}$$

As usual, absolutely continuous (AC) solutions to (1.1) and (2.4) are called, respectively, *ordinary trajectories* and *relaxed trajectories* to the original differential inclusion. For the proofs of our main results in the subsequent sections, we need the following facts concerning ordinary and relaxed trajectories to the differential inclusions under consideration, which are established in [8, 9] in more general settings.

Lemma 2.1. (Boundedness of Trajectories) *Let $x\colon T \to H$ be an absolutely continuous function satisfying the inclusion*

$$\dot{x}(t) \in \overline{\text{co}}\, F\big(t,x(t)+I\!B\big)+I\!B \ \textit{ for a.e. } \ t \in T, \quad x(0)=x_0 \tag{2.5}$$

under assumptions (A1) *and* (A2). *Then there is a number $M>0$ such that*

$$|x(t)| \le M \ \textit{ and } \ \sup\big\{|v| \ \big| \ v \in F\big(t,x(t)+I\!B\big)+I\!B\big\} \le M \ \textit{ for a.e. } \ t \in T. \tag{2.6}$$

Lemma 2.2. (Qualitative Properties of Solution Sets) *The following assertions hold under the standing assumptions* (A1) *and* (A2)*:*

(i) *The set of AC solutions to the relaxed differential inclusion* (2.4) *is nonempty and compact in the space $C(T;H)$ of continuous functions $x\colon T \to H$ endowed with the standard supremum norm on T.*

(ii) *Let $G\colon T \times H \rightrightarrows H$ be almost LSC with nonempty, compact values and such that*

$$G(t,x) \subset \overline{\text{co}}\, F(t,x) \ \textit{ for all } \ (t,x) \in T \times H.$$

Then the set of AC solutions to the differential inclusion

$$\dot{x}(t) \in G\big(t,x(t)\big) \ \textit{ for a.e. } \ t \in T, \quad x(0)=x_0 \tag{2.7}$$

is nonempty and $C(T;H)$-precompact, i.e., relatively compact in the norm topology of $C(T;H)$.

3 Strong Approximation of Solution Sets to MOSL Differential Inclusions under Discretization

The primary goal of this paper is to study *discrete approximations* to the original differential inclusion (1.1) satisfying the standing assumptions (A1) and (A2). For simplicity we consider the *uniform Euler scheme* to replace the time-derivative in (1.1) by the standard finite difference. Let

$$h := \frac{1}{k} \text{ and } t_j := jh, \quad j = 0, \dots, k, \quad k \in I\!N, \tag{3.1}$$

where we omit in notation the dependence on k of the discretization *stepsize* h and the *mesh* points t_j. The corresponding sequence of finite-difference inclusions is now given by

$$\begin{cases} z(t) = z(t_j) + (t - t_j)v_j, \quad z(0) = x_0, \quad t_j \le t \le t_{j+1}, \\ \text{with} \quad v_j \in F\big(t_j, z(t_j)\big), \quad j = 0, \dots, k-1, \end{cases} \tag{3.2}$$

where solutions to (3.2) are *piecewise linear* functions on T, i.e., they are familiar *Euler's polygons/broken lines.*

Due to the construction of (3.2), it is natural to expect that well-posedness and approximation results involving (3.2) require appropriate *continuity* assumptions on the dependence of F with respect to the *time* variable. One of the possibilities to avoid such requirements is to consider the sequence of *semi-discrete approximations* defined by

$$\begin{cases} \dot y(t) \in F(t, y(t_j)) \ \text{ a.e. } \ t \in [t_j, t_{j+1}), \quad y(t_j) := \lim\limits_{t \uparrow t_j} y(t), \\ j = 1, \dots, k-1, \quad y(0) = x_0, \end{cases} \tag{3.3}$$

which is well posed under the standing assumptions (A1) and (A2). In what follows, we denote by $\mathcal{S}$ the set of AC solutions to (1.1), by $\mathcal{S}(h)$ the set of AC solutions to (3.2) for any fixed h from (3.1), and by $\widetilde{\mathcal{S}}(h)$ the set of (absolutely continuous) solutions to (3.3).

In papers [7, 8, 9, 10, 11], the reader can find various estimates of the *uniform* Hausdorff distance—in the space $\mathcal{C}(T;H)$—between the solution set $\mathcal{S}$ to the *convex-valued* differential inclusion (1.1) and the solutions sets $\mathcal{S}(h)$ and $\widetilde{\mathcal{S}}(h)$ to its discretized counterparts under more general assumptions in comparison with (A1) and (A2). These results imply the *uniform convergence* of the sets $\mathcal{S}(h)$ and $\widetilde{\mathcal{S}}(h)$ to $\mathcal{S}$ as $h \downarrow 0$ in the space $\mathcal{C}(T;H)$; in particular, they imply the uniform approximation of solutions to (1.1) by solutions to the discretized inclusions (3.2) and (3.3). The latter corresponds, by the Newton-Leibnitz formula, to the *weak* convergence of the derivatives in $L^1(T;H)$.

Our main attention in this section is to obtain results on the *strong* in $L^1(T;H)$—actually in any $L^p(T;H)$ as $p \ge 1$ due to the assumptions made—convergence of the solution *derivatives* for sequences of the discrete and semi-discrete approximations, which implies the (almost everywhere) *pointwise convergence* of the corresponding subsequences. This means in fact the *strong convergence* of trajectories in the Sobolev spaces $W^{1,p}(T;H)$ instead of $\mathcal{C}(T;H)$ as before. Results of this type were derived in [16, 17, 18], for the case of discrete approximations (3.2) of differential inclusions with finite-dimensional and infinite-dimensional (reflexive) state spaces, under the *Lipschitz continuity* of F in x with no convexity assumptions on the velocity sets $F(t,x)$. In what follows we establish the strong convergence results, also for *nonconvex* inclusions while in the Hilbert space setting, under essentially less restrictive *MOSL* property of F. Such significant improvements of the previous results are important for their own sake and play a crucial role in applications to optimal control problems for MOSL differential inclusions considered in Section 5.

We start with relationships between *solution derivatives* for the differential inclusion (1.1) and its *semi-discrete* approximations (3.3). Denote by $\mathcal{D}$ and $\widetilde{\mathcal{D}}(h)$ the sets of the time-derivatives for solutions to (1.1) and (3.3), respectively. The next theorem justifies the *strong convergence* of the Hausdorff distance between these sets in the space $L^p(T;H)$, i.e., the *two-sided closeness* of these sets as $h \downarrow 0$.

Theorem 3.1. (Strong Convergence of Semi-Discrete Approximations for Nonconvex MOSL Differential Inclusions) *Under the standing assumptions* (A1) *and* (A2) *we have the solution set convergence*

$$d_{L^p}\big(\widetilde{\mathcal{D}}(h),\mathcal{D}\big) \to 0 \quad \textit{as} \quad h \downarrow 0 \quad \textit{for all} \quad p \ge 1, \tag{3.4}$$

where the Hausdorff distance is taken in the corresponding space $L^p(T;H)$.

Proof. It is sufficient to justify the strong convergence result of the theorem for the case of $p=1$, which easily implies (3.4) for any $p>1$ due to the standing assumptions made.

Observe that, by Lemma 2.1, every solution $y(\cdot)$ to (3.3) for all $h>0$ sufficiently small— which is always assumed in what follows— can be extended to the whole interval T, and we have the estimate

$$\sup\big\{|v| \;\big|\; v \in F\big(t,y(t)+I\!B\big)+I\!B\big\} \le M \quad \text{for a.e.} \quad t \in T. \tag{3.5}$$

Hence the sets $\widetilde{\mathcal{D}}(h)$ as $h>0$ are *uniformly bounded* in $L^1(T;H)$ together with the sets $\mathcal{S}$ and $\mathcal{D}$, which are *nonempty* by Lemma 2.2. Furthermore, the sets $\mathcal{D}$ and $\widetilde{\mathcal{D}}(h)$ are obviously closed in the norm topology of $L^1(T;H)$, and thus the Hausdorff distance $d_{L^1}\big(\widetilde{\mathcal{D}}(h),\mathcal{D}\big)$ between them is well defined.

Part 1. We first prove that the set $\mathcal{D}$ can be approximated by $\widetilde{\mathcal{D}}(h)$ as $h \downarrow 0$ in the space $L^1(T;H)$. Take any $x(\cdot) \in \mathcal{S}$ and construct the required discrete approximations $y(\cdot) \in \widetilde{\mathcal{S}}(h)$ as $h \downarrow 0$ of this trajectory by the following *step-by-step* procedure on the consequent intervals $[t_j,t_{j+1}]$ for $j=0,\ldots,k-1$. Denote $y_j := y(t_j)$ for $j=0,\ldots,k-1$ and observe that—since the initial point $y_0 = x_0$ is given—it is sufficient to construct the required trajectory $y(t)$ to (3.3) on the interval $[t_j,t_{j+1}]$ for $j=0,\ldots,k-1$ *provided* that $y_j = y(t_j)$ is known. To proceed, let us show that whenever $j=0,\ldots,k-1$ there is a *strongly measurable selection*

$$v_j(t) \in F(t,y_j) \quad \text{for a.e.} \quad t \in [t_j,t_{j+1}] \tag{3.6}$$

satisfying the relationships

$$\langle y_j - x(t), v_j(t) - \dot{x}(t)\rangle \le L|y_j - x(t)|^2 \quad \text{and} \quad |v_j(t)-\dot{x}(t)| \le f\big(t,|x(t)-y_j|\big) \tag{3.7}$$

for a.e. $t \in [t_j,t_{j+1}]$ as $j=0,\ldots,k-1$.

Indeed, it is easy to check that for each $j = 0, \ldots, k-1$ the set-valued mapping $S_j \colon [t_j, t_{j+1}] \rightrightarrows H$ defined by

$$S_j(t) := \big\{v \in H \big| \, \langle y_j - x(t), v(t) - \dot{x}(t)\rangle \le L|y_j - x(t)|^2, \quad |\dot{x}(t) - v| \le f\big(t, |x(t) - y_j|\big)\big\}$$

is *Lusin* in the sense of [5], and hence it is measurable on $[t_j, t_{j+1}]$. Consequently, each intersection mapping Q_j defined by

$$Q_j(t) := F(t, y_j) \cap S_j(t), \quad t \in [t_j, t_{j+1}], \quad j = 0, \ldots, k-1,$$

is *nonempty-valued* due to (A2) and *strongly measurable* on $[t_j, t_{j+1}]$, since $F(t, y_j)$ is compact-valued and has this property by (A1). Thus, by the classical *measurable selection* results and the *almost separable-valuedness* of Q_j (see, e.g., [20, Chapter 1]), there is a strongly measurable selection $v_j(t) \in Q_j(t)$ for a.e. $t \in [t_j, t_{j+1}]$ satisfying the relationships in (3.6) and (3.7) whenever $j = 0, \ldots, k-1$. Moreover, each selection $v_j(\cdot)$ is actually *summable* on the corresponding interval $[t_j, t_{j+1}]$ by the boundedness property (3.5).

Having in hand the solution $x(t)$ to (1.1) and the summable selections $v_j(t)$ satisfying (3.6) and (3.7) for a.e. $t \in [t_j, t_{j+1}]$ with $j = 0, \ldots, k-1$, we construct the corresponding solution $y(t)$ to (3.3) defining it on each interval $[t_j, t_{j+1}]$ by

$$y(t) := y_j + \int_{t_j}^{t} v_j(s)\, ds \quad \text{for all} \quad t \in [t_j, t_{j+1}], \quad j = 0, \ldots, k-1, \tag{3.8}$$

where the integral is taken in the *Bochner sense*, and thus $y(\cdot)$ satisfies the differential inclusion (3.3). Furthermore, by (3.5) and (3.7), we have the following estimates for a.e. $t \in [t_j, t_{j+1}]$ and all $j = 0, \ldots, k-1$:

$$\begin{aligned}
&\langle y(t) - x(t), v_j(t) - \dot{x}(t)\rangle \le L|y(t) - x(t)|^2 \\
&\le |L|\big(|y(t) - x(t)|^2 - |y_j - x(t)|^2\big) + |v_j(t) - \dot{x}(t)| \cdot |y_j - y(t)| \\
&\le L|y(t) - x(t)|^2 + 2M^2\big(2|L| + 1\big)h.
\end{aligned}$$

This consequently implies the inequalities

$$\frac{d}{dt}|y(t) - x(t)|^2 \le 2L|y(t) - x(t)|^2 + Ch, \quad |y(t) - x(t)| \le C\sqrt{h}$$

and thus gives by (3.7) the desired estimate

$$|\dot{y}(t) - \dot{x}(t)| \le f(t, C\sqrt{h}) \quad \text{for a.e.} \quad t \in T,$$

where $C > 0$ is a generic constant. By the properties of f in (A2) we therefore get the strong $L^1(T; H)$-convergence of $y(\cdot) = y_h(\cdot)$ to $x(\cdot)$ as $h \downarrow 0$ and finish the proof of Part 1.

Part 2. Let us now show that, taking any solution $y(\cdot)\widetilde{S}(h)$ to the semi-discrete inclusion (3.3), we always can find a solution $x(\cdot) \in S$ to the original differential inclusion (1.1) such that

$$|\dot{x}(t) - \dot{y}(t)| \le f(t, C\sqrt{h}) \quad \text{for a.e.} \quad t \in T, \tag{3.9}$$

where f is our standing estimate function from (A2) while $C > 0$ is a generic constant. It is clear that estimate (3.9) implies the required approximation of the derivative set $\widetilde{\mathcal{D}}(h)$ for (3.3) by the derivative set $\mathcal{D}$ for (1.1), and hence—together with Part 1—it fully justifies the claimed convergence (3.4) of the theorem.

To proceed with the proof of (3.9), we take any $\varepsilon > 0$ and consider the set-valued mapping $G_\varepsilon \colon T \times H \rightrightarrows H$ defined by

$$G_\varepsilon(t,x) := \overline{\big\{v \in F(t,x)\big|\ \langle y_j - x, \dot{y}(t) - v\rangle < L|y_j - x|^2 + \varepsilon,\ \ |\dot{y}(t) - v| < f(t,|y_j - x|) + \varepsilon\big\}}$$

for a.e. $t \in [t_j, t_{j+1}]$ with $y_j = y(t_j)$ as $j = 0, \ldots, k-1$. Since the original mapping F is compact-valued, so is G_ε, and—due to the basic assumption (A2)—the values of $G_\varepsilon(t,x)$ are *nonempty* for all $x \in H$ and a.e. $t \in T$. Moreover, it is standard to check that the constructed mapping G_ε is *almost LSC* for any $\varepsilon > 0$. Employing now Lemma 2.2(ii), we conclude that the differential inclusion

$$\dot{x}(t) \in G_\varepsilon\big(t, x(t)\big), \quad x(0) = x_0 \tag{3.10}$$

admits an AC solution $x(\cdot)$ on T. It further follows from the construction of G_ε—by the MOSL property of F—that

$$\frac{d}{dt}\big|x(t) - y(t)\big|^2 < 2L|x(t) - y(t)|^2 + C(h + \varepsilon) \ \text{ for a.e. } \ t \in T,$$

which consequently implies the inequalities

$$\big|x(t) - y(t)\big| < C\sqrt{h + \varepsilon} \ \text{ on } \ T \ \text{ and } \ \big|\dot{x}(t) - \dot{y}(t)\big| \le f\big(t, C\sqrt{h+\varepsilon}\big) \ \text{ for a.e. } \ t \in T.$$

Since $\varepsilon > 0$ was chosen arbitrarily, we arrive at the required estimate (3.9) and thus complete the proof of Part 2 and of the whole theorem. □

Next we study the *strong approximation*—in the norm topology of $W^{1,p}(T; H)$—of *any feasible* trajectory $x(\cdot) \in \mathcal{S}$ to the original *nonconvex* differential inclusion (1.1) satisfying the *MOSL condition* by a sequence of feasible trajectories $z_k(\cdot) \in \mathcal{S}(h_k)$ to the *discrete inclusions* (3.2). We establish *two independent versions* of such a strong approximation result. The fist version presented in what follows justifies the strong discrete approximation for *any sequence* of partitions of the interval T—even for nonuniform partitions more general than (3.1)—imposing, however, additional *continuity* assumptions on the mappings F and f with respect to *both* variables (t, x). The second version drops these additional assumptions and imposes only the standing assumptions (A1) and (A2), but the price to pay is that the strong convergence can be justified only for *some sequence* of discrete partitions of T. Since the proof of the second version is technically more involved and is strongly based on the technique developed in the proof of the density theorem in Section 4, it makes sense to present the latter version in the next section.

Theorem 3.2. (Strong Convergence of Discrete Approximations for Nonconvex MOSL Differential Inclusions under Continuity Assumptions) *Suppose that the* *mappings* F, f *in assumptions* (A1) *and* (A2) *are continuous in both variables. Then for every AC solution* $x(\cdot)$ *to* (1.1) *and for every sequence of partitions* Δ_k *of* T *given by*

$$\Delta_k := \{0 = t_0^k < t_1^k < \ldots < t_k^k = 1\} \quad \text{with} \quad h_k := \max_{0\le j\le k-1} \{t_{j+1}^k - t_j^k\} \downarrow 0 \tag{3.11}$$

there is a sequence of piecewise linear solutions $z_k(\cdot)$ *to the discretized inclusions* (3.2) *on* Δ_k *satisfying the relationships*

$$z_k(t) \to x(t) \quad \text{uniformly on } T \quad \text{and} \quad \int_0^1 \big|\dot z_k(t) - \dot x(t)\big|^p\,dt \to 0, \quad p \ge 1, \tag{3.12}$$

as $k \to \infty$*; the latter implies the convergence* $\dot z_k(t) \to \dot x(t)$ *of a subsequence for a.e.* $t \in T$.

Proof. Fix an arbitrary number $\varepsilon > 0$ and observe—by Lemma 2.1—that there is a constant $M > 0$ such that for a.e. $t \in T$ we have the estimate

$$|x(t)| \le M \quad \text{and} \quad |\dot x(t)| \le M \quad \text{whenever} \quad \operatorname{dist}\big(\dot x(t), F\big(t, x(t)\big)\big) < \varepsilon. \tag{3.13}$$

Note also that, due to the continuity of $F(\cdot,\cdot)$, the composition $F\big(t, x(t)\big)$ is *uniformly continuous* (in on the compact interval T for any continuous function $x\colon T \to H$.

To proceed, we pick an AC solution $x(\cdot)$ to (1.1) with the derivative $\dot x(t)$ and consider the given sequence of partitions Δ_k from (3.11). As mentioned, it is sufficient to justify (3.12) for $p = 1$. By the *density* of step functions in $L^1(T;H)$, approximate $\dot x(t)$ *strongly* in $L^1(T;H)$ by a sequence of step functions $w_k(t)$, which are *bounded* in $L^1(T;H)$ and *constant* on the intervals $[t_j^k, t_{j+1}^k)$, $j = 0, \ldots, k-1$, from the sequence of partitions (3.11). The latter can be adopted without loss of generality in the proof below due to the continuity assumptions imposed. Construct now the AC functions

$$y_k(t) := x_0 + \int_0^t w_k(s)\,ds, \quad t \in T, \quad k \in I\!N, \tag{3.14}$$

via the Bochner integral of $w_k(\cdot)$ and observe that

$$y_k(t) \to x(t) \quad \text{uniformly in } t \in T \text{ as } k \to \infty.$$

Since $w_k(t) \to \dot x(t)$ *pointwisely* on T along a subsequence of $k \to \infty$ and since $w_k(\cdot)$ are piecewise constant, we can select $\widetilde t_j^k \in [t_j^k, t_{j+1}^k)$ such that

$$\big|w_k(\widetilde t_j^k) - \dot x(\widetilde t_j^k)\big| \le \varepsilon/2 \quad \text{for all } j = 0, \ldots, k-1 \text{ and } k \in I\!N \tag{3.15}$$

and that the differential inclusion (1.1) holds at $t = \widetilde t_j^k$.

Let us show next that

$$\operatorname{dist}\big(w_k(t); F(t, y_k(t)\big) \le \varepsilon \quad \text{whenever} \quad t \in T \tag{3.16}$$

and $k \in I\!N$ is sufficiently large; in the latter case we include *all* $k \in I\!N$ into consideration. Indeed, select $k \in I\!N$ so large that

$$d_H\big(F(t, x(t)), F(\widetilde t_j^k, x(\widetilde t_j^k))\big) \le \varepsilon/2 \quad \text{for all } j = 0, \ldots, k-1 \text{ and such } k \in I\!N,$$

for the Hausdorff distance in H, which is possible due to the choice of $\widetilde{t}_j^k$ and the uniform continuity of $F(t,x(t))$ on T. Then using again the continuity of $F(\cdot,\cdot)$ and the uniform convergence of $y_k(\cdot) \to x(\cdot)$ on T, we get

$$\operatorname{dist}\big(w_k(t); F(t,y_k(t))\big) \le \operatorname{dist}\big(w_k(\widetilde{t}_j^k); F(\widetilde{t}_j^k, x(\widetilde{t}_j^k))\big) + d_H\big(F(t,y_k(t)), F(\widetilde{t}_j^k, x(\widetilde{t}_j^k))\big)$$

for all $t \in T$ and for all large k, since $w_k(\cdot)$ are piecewise constant on $[t_j^k, t_{j+1}^k)$ and satisfy (3.15). This justifies (3.16).

Observe that the functions $y_k(\cdot)$ defined in (3.14) are *not* feasible trajectories to the discretized inclusions (3.2). Now we construct, based on $y_k(\cdot)$ and the *MOSL property* of F in (A2), the required piecewise trajectories $z_k(\cdot)$ to inclusions (3.2) on the partitions Δ_k built above such that the *strong* convergence relationships (3.12) are satisfied.

Fix $k \in I\!N$ and construct the required trajectory $z(t) = z_k(t)$ to (3.2) on Δ_k omitting the index "k" in the notation of $z(t)$ and $t_j = t_j^k$ for simplicity. We proceed as follows. Assuming that $z(t_j)$ is known (for $j = 0$ it is always the case), we want to extend $z(\cdot)$ to the interval $(t_j, t_{j+1}]$ in (3.2). By the structure of (3.2) this means that we need to find an appropriate velocity $v_j \in F\big(t_j, z(t_j)\big)$. Let us do it by the *projection method* on the base of the MOSL property of $F(t_j,\cdot)$. Having $w_k(t_j)$ and $y_k(t_j)$ from the above constructions, we select—by the compactness of $F(t,x)$—a Euclidean projection

$$u_j \in \operatorname{proj}_{w_k(t_j)} F\big(t_j, y_k(t_j)\big)$$

for this fixed $j \in \{0,\ldots,k-1\}$. Note that $|u_j| \le M$ and $|u_j - w_k(t_j)| \le \varepsilon$ by (3.13) and (3.15). Employing the MOSL property (A2) of $F(t_j,\cdot)$ with $x_1 = y_k(t_j)$, $x_2 = z(t_j)$, and $u_j \in F\big(t_j, y_k(t_j)\big)$, we find $v_j \in F\big(t_j, z(t_j)\big)$ satisfying

$$\langle y_k(t_j) - z(t_j), u_j - v_j\rangle \le L|y_k(t_j) - z(t_j)|^2, \quad |u_j - v_j| \le f\big(t_j, |y_k(t_j) - z(t_j)|\big).$$

Define now the trajectory $y(t)$ of (3.2) on $[t_j, t_{j+1}]$ by using this velocity v_j and show that the constructed sequence $z_k(t) = z(t)$ on T satisfies the required properties. By the choice of v_j and the triangle inequality we have

$$\begin{aligned}
&\langle y_k(t) - z(t), u_j - v_j\rangle \le \langle y_k(t_j) - z(t_j), u_j - v_j\rangle\\
&+|\langle y_k(t) - z(t), u_j - v_j\rangle - \langle y_k(t_j) - z(t_j), u_j - v_j\rangle|\\
&\le L|y_k(t_j) - z(t_j)|^2 + \big(|u_j| + |v_j|\big)\big(|z(t) - z(t_j)| + |y_k(t) - y_k(t_j)|\big)\\
&\le L|y_k(t) - z(t)|^2 + \big|L|y_k(t) - z(t)|^2 - L|y_k(t_j) - z(t_j)|^2\big| + 4M^2(t - t_j).
\end{aligned}$$

The latter implies by elementary transformations that

$$\begin{aligned}
&\big|\,|y_k(t) - z(t)|^2 - |y_k(t_j) - z(t_j)|^2\big|\\
&\le \big(|y_k(t)| + |z(t)| + |y_k(t_j)| + |z(t_j)|\big)\big(|y_k(t) - z(t_j)| + |z(t) - z(t_j)|\big)\\
&\le 8M^2 h_k \ \text{ for all } \ t \in [t_j, t_{j+1}], \quad j = 0,\ldots,k-1, \quad k \in I\!N.
\end{aligned}$$

Furthermore, taking into into account that $|u_j - v_j| \le \varepsilon$ for all $j = 0,\ldots,k-1$ by the above constructions of u_j, v_j and the previous estimates, we get

$$\frac{d}{dt}\big|y_k(t) - z(t)\big|^2 \le 2L|y_k(t) - z(t)|^2 + C(h_k + \varepsilon) \ \text{ for a.e. } \ t \in [t_j, t_{j+1}], \quad j = 0,\ldots,k-1,$$

with a generic constant $C > 0$, which consequently gives

$$|y_k(t) - z(t)| \le C\sqrt{h_k + \varepsilon} \text{ and } |\dot{y}_k(t) - \dot{z}(t)| \le \varepsilon + \max_{t \in T} f\big(t, \sqrt{h_k + \varepsilon}\big), \quad t \in T, \quad (3.17)$$

for all $k \in I\!N$. Putting $\varepsilon = \varepsilon_k \downarrow 0$ as $k \to \infty$ in (3.17) and using the uniform convergence of $y_k(t) \to x(t)$ as well as the L^p-convergence of $\dot{y}_k(\cdot) = w_k(\cdot) \to \dot{x}(\cdot)$, we arrive at the claimed relationships (3.12) and complete the proof of the theorem. □

4 Bogolyubov-Type Relaxation Theorem for MOSL Differential Inclusions

This section concerns relationships between the original dynamic system (1.1) and its convexification (2.4). Questions of this type play a key role in many aspects of dynamic optimization and related topics; they are usually unified under the name of *density* and/or *relaxation* theorems. In the framework of the classical calculus of variations, pioneering research was done by Bogolyubov, Young, and McShane in the 1930s; in optimal control–by Gamkrelidze, Filippov, Warga, and Ważewski in the 1960s. The reader can find more information and discussions, e.g., in the books [2, 13, 17, 20] and the references therein.

Relaxation/density results say, roughly speaking, that admissible trajectories to the original continuous-time dynamic system are *dense* under certain conditions among admissible trajectories to the convexified/relaxed one and, furthermore, that the value of the cost functional in the corresponding dynamic optimization problem does *not change* under convexification. The first result of this type was probably obtained by Bogolyubov [3] for the simplest problem of the calculus of variations; and thus results in this vein are often called Bogolyubov-type theorems.

We refer the reader to [2, 4, 20] for the classical and recent results in this direction for differential inclusions in finite-dimensional and infinite-dimensional spaces. These results are obtained under the *full Lipschitz* condition imposed on the velocity map F with respect to the state variable. Moreover, the classical example by Plis [19] (see also [20, Example 3.2.1]) shows that the Lipschitz continuity of $F(t, \cdot)$ *cannot* be dropped, or even relaxed to continuity. In fact, Plis' example corresponds to system (2.3) with the *only change*: the term $-\text{sign}(x_2)$ is replaced with $\text{sign}(x_2)$. As mentioned above in Section 2, density/relaxation results do *not* generally hold if the Lipschitz continuity of $F(t, \cdot)$ is replaced with its *one-sided* Lipschitz continuity.

The primary goal of this section is to show that the *modified* one-sided Lipschitz condition (A2) allows us to establish appropriate density/relation results, which are further employed in Section 5 to the strong convergence of discrete approximations. Note, in particular, that the "almost-Plis" system (2.3) satisfies our requirements.

To cover in the sequel dynamic optimization problems of the Bolza type, we consider—along with the original differential inclusion (1.1)—the integral functional

$$I[x] := \int_0^1 g\big(t, x(t), \dot{x}(t)\big)\, dt \quad (4.1)$$

defined over absolutely continuous trajectories $x\colon T \to H$ ($T = [0,1]$) to the differential inclusion (1.1). In addition to the standing assumptions (A1) and (A2) on $F(t,x)$, we impose the following ones on the integrand g in the "cost" functional in (4.1):

(A3) The integrand $g\colon T \times H \times H \to I\!R$ in (4.1) is almost continuous on the product $T \times (H \times H)$ and its absolute value is majorized by a summable function on T uniformly in the last two variables.

Note that the uniform boundedness assumptions on the integrand g is imposed for simplicity; it can be replaced by an appropriate *growth condition* as, e.g., in [4].

Consider further the following *extended* differential system involving the differential inclusion (1.1) and the differential equation generated by (4.1):

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big) \ \text{ for a.e. } \ t \in T, \quad x(0) = x_0, \\ \dot{s}(t) = g\big(t, x(t), \dot{x}(t)\big) \ \text{ for a.e. } \ t \in t, \quad s(0) = 0. \end{cases} \tag{4.2}$$

Letting $y := (x, s) \in H \times I\!R$, define the set-valued mapping $G\colon T \times H \times I\!R \rightrightarrows H \times I\!R$ by

$$G(t,y) := \big\{(v,\vartheta) \in H \times I\!R \big|\ v \in F(t,x);\ \vartheta = g(t,x,v)\big\}, \tag{4.3}$$

and consider the *extended differential inclusion*

$$\dot{y}(t) \in G(t,y) \ \text{ for a.e. } \ t \in T, \quad y(0) = y_0 := (x_0, 0) \tag{4.4}$$

together with its *relaxation/convexification*

$$\dot{y}(t) \in \overline{\text{co}}\, G(t,y) \ \text{ for a.e. } \ t \in T, \quad y(0) = y_0. \tag{4.5}$$

Observe that the extended differential inclusion (4.4) is obviously *equivalent* to the extended system (4.2) and that the mapping G in (4.4) is actually *independent* of the component $s \in I\!R$ of the state variable $y = (x,s)$. The following new *density theorem* establishes the possibility of the *uniform* approximation—under the *key MOSL condition*—of any AC trajectory to the convexified extended inclusion (4.5) by AC trajectories to its ordinary counterpart (4.4).

Theorem 4.1. (Uniform Density under Relaxation of MOSL Differential Inclusions) *Let all the assumptions* (A1), (A2), *and* (A3) *be satisfied. The the set of AC solutions to the extended differential inclusion* (4.4) *is dense with respect to the norm topology of $C(T;H)$ in the set of AC solutions to the convexified differential inclusion* (4.5).

Proof. It is easy to observe that the mapping $G(\cdot,\cdot)$ in (4.3) is *almost continuous* due to imposing this property on F and g. Furthermore, we conclude from the boundedness assumptions in (A1) and (A3) and the boundedness property of Lemma 2.1 that the sets $G(t,y) = G(t,x)$ are *uniformly bounded* over a bounded set containing all the relaxed trajectories. For definiteness, suppose that

$$\sup\big\{|u| \ \big|\ u \in \overline{\text{co}}\, G(t,y)\big\} \le M - 1/2 \ \text{ with some } \ M > 1/2 \tag{4.6}$$

for all (t,y) under consideration. Let us now fix an arbitrary AC trajectory $z(t)$ to the convexified inclusion (4.5). Our goal is, given any $\varepsilon > 0$, to ε-approximate it in the norm topology of $\mathcal{C}(T;H)$ by an AC trajectory to the extended differential inclusion (4.4). We split our proof into *two major steps*; each of them is certainly if independent interest.

Step 1. First we find a *quasitrajectory* $w(\cdot)$ to (4.4), which is ε-close to $z(\cdot)$ in the norm of $\mathcal{C}(T;H)$. Our intention thus is to construct an AC function $w\colon T \to H$ such that

$$\dot{w}(t) \in G\big(t, w(t)\big) + \varepsilon I\!\!B \quad \text{as} \quad t \in T_\varepsilon, \quad w(0) = y_0 \tag{4.7}$$

for some compact subset $T_\varepsilon \subset T$ with $\text{mes}(T_\varepsilon) > 1 - \varepsilon$, that $|\dot{w}(t)| \le M$ on T_ε, and that

$$\big|w(t) - z(t)\big| \le \varepsilon \quad \text{for all} \quad t \in T. \tag{4.8}$$

Taking (4.6) into account, we have from (4.7) and (4.8) that

$$\text{dist}\big(\dot{w}(t); G(t, w(t))\big) \le 2M \quad \text{on} \quad T \setminus T_\varepsilon,$$

which we use in what follows. Note that in the proof of Step 1 below we do *not* employ the MOSL property of F while manage to establish the approximation result by quasitrajectories under merely the *almost continuity* assumption on F and g, which are *weaker* than in previously known results of this type in both finite and infinite dimensions; see, e.g., [2, 4, 20] and the references therein.

To begin with, take $\lambda > 0$ and show that there exist a compact set $T_\lambda \subset T$ with $\text{mes}(T_\lambda) > 1 - \lambda^2$ and an absolutely continuous function $p\colon T \to H$ with the piecewise constant derivative satisfying

$$\|\dot{z} - \dot{p}\|_{L^1(T;H)} \le \lambda \quad \text{and} \quad \text{dist}\big(\dot{p}(t); \overline{\text{co}}\, G(t, p(t))\big) \le \lambda/10 \quad \text{on} \quad T_\lambda. \tag{4.9}$$

Indeed, by the almost continuity property of $G(\cdot,\cdot)$ and the classical Lusin property of $\dot{z}(\cdot)$, we find $T_\lambda \subset T$ with $\text{mes}(T_\lambda) > 1 - \lambda^2$ such that $G(\cdot,\cdot)$ is continuous on $T_\lambda \times H$ and that $\dot{z}(\cdot)$ is continuous on T_λ. Since the convexified mapping $\overline{\text{co}}\, G(\cdot,\cdot)$ is also continuous on $T_\lambda \times H$, for some $\gamma \in (0, \lambda/20)$ we have

$$d_H\big(G(t, z(t)), G(t, y)\big) \le \lambda/20 \quad \text{and} \quad d_H\big(\overline{\text{co}}\, G(t, z(t)), \overline{\text{co}}\, G(t, y)\big) \le \lambda/20 \tag{4.10}$$

whenever $|z(t) - y| \le \gamma$ and $t \in T_\lambda$. Employing the classical Egorov theorem from real analysis and taking into account that $\dot{z}(t)$ is uniformly continuous on the compact set T_λ, find a piecewise constant function $v\colon T \to H$ such that

$$|\dot{z}(t) - v(t)| \le \gamma/20 \quad \text{for} \quad t \in T_\lambda \quad \text{and} \quad \|\dot{z} - v\|_{L^1(T;H)} \le \gamma.$$

Defining now $p(\cdot)$ by the Bochner integral

$$p(t) := y_0 + \int_0^t v(\tau)\, d\tau, \quad t \in T,$$

and taking into account the choice of $\gamma > 0$, we get the desired function $p(\cdot)$ satisfying the relationships in (4.9) . Clearly, $|z(t) - p(t)| \le \gamma$ on T.

Having this function in hand, let us construct the *approximating quasitrajectory* $w(\cdot)$ to (4.4) with the properties described above. To proceed, we divide the underlying interval $T = [0,1]$ into nonintersecting and depending on the chosen $\lambda > 0$ intervals $\{J_k\}$, $k \in I\!N$, with lengths not greater than λ^2 such that the integrand $v(\cdot)$ is constant on each J_k and

$$d_H\big(G(t,p(t)),G(\tau,p(\tau))\big) \le \lambda/10 \ \text{ whenever } \ t,\tau \in J_k \cap T_\lambda, \quad k \in I\!N. \tag{4.11}$$

Take some $\tau_k \in J_k \cap T_\lambda$ for each $k \in I\!N$ and consider the *projection*

$$\pi_k := \text{proj}_{\,v(\tau_k)} \overline{\text{co}}\, G\big(\tau_k, p_k(\tau_k)\big)$$

of the point $v(\tau_k)$ on the set $\overline{\text{co}}\, G\big(\tau_k, p_k(\tau_k)\big)$; the existence and uniqueness of this projection under the assumptions made are well known. By (4.10) we obviously have the estimate

$$|\pi_k - v(\tau_k)| \le \lambda/10 \ \text{ for all } \ k \in I\!N. \tag{4.12}$$

Consequently, there are $\alpha_k^i \ge 0$ and $u_k^i \in G\big(\tau_k, p(\tau_k)\big)$ for $i = 1,\ldots,m_k$ with some $m_k \in I\!N$ such that, by taking the closure operation in (4.12) into account, we get the relationships

$$\sum_{i=1}^{m_k} \alpha_k^i = 1 \ \text{ and } \ \Big|\pi_k - \sum_{i=1}^{m_k} \alpha_k^i u_k^i\Big| \le \frac{\lambda}{10}$$

whenever $m_k \in I\!N$ is sufficiently large. For every fixed $k \in I\!N$ we divide now the interval J_k into m_k *pairwise disjoint* measurable sets $\mathcal{J}_k^i$ such that

$$\alpha_k^i = \frac{\text{mes}(\mathcal{J}_k^i)}{\text{mes}(J_k)}, \quad i = 1,\ldots,m_k, \quad k \in I\!N.$$

Since the union of J_k over $k \in I\!N$ gives the whole interval T, and the union of the sets $\mathcal{J}_k^i$ over $i \in \{1,\ldots,m_k\}$ gives J_k for each k, we can construct—for the chosen $\lambda > 0$—the summable function $u_\lambda\colon T \to H$ by

$$u_\lambda(t) := u_k^i \ \text{ for } \ t \in \mathcal{J}_k^i, \quad i = 1,\ldots,m_k, \quad k \in I\!N,$$

and then define the absolutely continuous function $w_\lambda\colon T \to H$ by the Bochner integral

$$w_\lambda(t) := y_0 + \int_0^t u_\lambda(\tau)\,d\tau, \quad t \in T. \tag{4.13}$$

It is easy to observe from the above estimates that

$$|w_\lambda(t) - p(t)| \le \lambda/4 \ \text{ and } \ |w_\lambda(t) - z(t)| \le \lambda/2 \ \text{ for all } \ t \in T.$$

Finally, we select $\lambda = \lambda(\varepsilon) < \varepsilon$ so small that

$$d_H\big(G(t,z(t)),G(t,y))\big) \le \varepsilon/3 \ \text{ on } \ T_\varepsilon \subseteq T_\lambda \ \text{ whenever } \ |z(t) - y| \le \lambda. \tag{4.14}$$

The latter estimate and the inequality (4.11) imply the following estimates for the function $w(t) = w_{\lambda(\varepsilon)}(t)$ constructed in (4.13):

$$\begin{aligned} \text{dist}\big(\dot w(t); G(t,w(t))\big) &\le \text{dist}\big(\dot w(t); G(t,p(t))\big) + d_H\big(G(t,p(t)),G(t,w(t))\big) \\ &\le d_H\big(G(\tau_k,p(\tau_k)),G(t,p(t))\big) + d_H\big(G(t,p(t)),G(t,w(t))\big). \end{aligned}$$

The triangle inequality

$$d_H\big(G(t,p(t)),G(t,w(t))\big) \le d_H\big(G(t,p(t)),G(t,z(t))\big) + d_H\big(G(t,z(t)),G(t,w(t))\big)$$

together with (4.11) and (4.14) imply (4.7) on T_ε, which shows that this function $w(t)$ is the required quasitrajectory to (4.4) satisfying the relationships in (4.7) and (4.8). This completes the proof of Step 1.

Step 2. Next we are going to show that the *quasitrajectory* $w(\cdot)$ to (4.4) constructed above can be approximated by a proper AC *trajectory* $y(\cdot)$ to this differential inclusion. To accomplish this goal, we strongly use the *MOSL property* of the original velocity mapping F, which turns out to be a crucial assumption replacing the full Lipschitz continuity in both finite and infinite dimensions. Having $w(t)$ that satisfies (4.7) and (4.8), we represent it as $w(t) = \big(q(t),\vartheta(t)\big)$ with $q\colon T \to H$ and $\vartheta\colon T \to I\!R$; clearly the q-part of w satisfies the differential inclusion

$$\dot q(t) \in F\big(t,q(t)\big) + \varepsilon I\!B \ \text{ on } \ T_\varepsilon, \quad q(0) = x_0, \tag{4.15}$$

where the compact $T_\varepsilon \subset T$ is described in the beginning of Step 1. By using the compactness of the velocity sets $F(t,x)$ and *measurable selection* theorems (cf. the proof of Theorem 3.1), we can select the *projection*

$$\pi(t) \in \text{proj}_{\,\dot q(t)} F\big(t,q(t)\big) \ \text{ on } \ T_\varepsilon,$$

which is *strongly measurable* on this set. Further, fix $\gamma > 0$ and define the multifunction

$$\begin{aligned} P_\gamma(t,u) := \Big\{ \ v \in F(t,u)\Big| \ & |v - \pi(t)| < f\big(t,|q(t) - u|\big) + \varepsilon + \gamma, \\ & \langle q(t) - u, \pi(t) - v\rangle < L|q(t) - u|^2 + \varepsilon + \gamma\Big\}, \quad t \in T_\varepsilon, \end{aligned}$$

where the constant $L \in I\!R$ and the function $f\colon T \times I\!R_+ \to I\!R_+$ are taken from the MOSL assumption (A2). Denote

$$Q_\gamma(t,u) := \begin{cases} \overline{P_\gamma(t,u)} & \text{if } \ t \in T_\varepsilon, \\ F(t,u) & \text{otherwise} \end{cases} \tag{4.16}$$

and observe that $Q_\gamma(\cdot,\cdot)$ has *nonempty* and compact values due to (A2). Let us show now that this mapping is *almost LSC*.

Since $F(\cdot,\cdot)$ is almost continuous and $\pi(\cdot)$ is measurable, for any $\nu > 0$ we find a compact set $T_\nu \subset T$ with $\text{mes}(T \setminus T_\nu) < \nu$ such that $F(\cdot,\cdot)$ is continuous on $T_\nu \times H$ and $\pi(\cdot)$ is continuous on T_ν. Then it easily follows from the construction of $Q_\gamma(\cdot,\cdot)$ in (4.16) that this mapping is LSC on $T_\nu \times H$, and so it is almost LSC on $T \times H$. Applying now Lemma 2.2(ii), we conclude that there is an AC function $q_\gamma\colon T \to H$ satisfying the differential inclusion

$$\dot q_\gamma(t) \in Q_\gamma\big(t,q_\gamma(t)\big) \ \text{ for a.e. } \ t \in T, \quad q_\gamma(0) = x_0. \tag{4.17}$$

It easily follows from (4.15)–(4.17) that

$$\begin{aligned} &\langle q(t) - q_\gamma(t), \dot q(t) - \dot q_\gamma(t)\rangle < L|q(t) - q_\gamma(t)|^2 + |\langle q(t) - q_\gamma(t), \dot q(t) - \pi(t)\rangle + \varepsilon + \gamma \\ &\le L|q(t) - q_\gamma(t)|^2 + \varepsilon|q(t) - q_\gamma(t)| + \varepsilon + \gamma \\ &\le L|q(t) - q_\gamma(t)|^2 + 0.5\big(\varepsilon^2 + |q(t) - q_\gamma(t)|^2\big) + \varepsilon + \gamma, \quad t \in T_\varepsilon. \end{aligned}$$

This consequently implies the estimate

$$|q(t) - q_\gamma(t)|^2 \le r(t) \text{ for all } t \in T, \tag{4.18}$$

where $r(0) = 0$ and the absolutely continuous function $r: T \to I\!R_+$ satisfies the following *differential inequalities* on T_ε and $T \setminus T_\varepsilon$, respectively:

$$\dot{r}(t) \le \begin{cases} 2(L+1)r(t) + \varepsilon^2 + 2(\varepsilon + \gamma) & \text{on } T_\varepsilon, \\ r(t) + 16M^2 & \text{on } T \setminus T_\varepsilon. \end{cases} \tag{4.19}$$

Applying the classical *Gronwall Lemma* to (4.19), we get the estimate

$$r(t) \le C(\varepsilon + \gamma) \text{ for all } t \in T, \tag{4.20}$$

where the generic constant C is independent of ε and γ. Thus

$$|q(t) - q_\gamma(t)| \le C\sqrt{\varepsilon + \gamma} \text{ for all } t \in T \tag{4.21}$$

by (4.18) and (4.20) Consider now the integral functional (4.1), with a variable upper limit of integration $t \in T$, computed on the absolutely continuous functions $q: T \to H$ and $q_\gamma: T \to H$, respectively:

$$\vartheta(t) = \int_0^t g\big(\tau, q(\tau), \dot{q}(\tau)\big)\, d\tau, \quad \vartheta_\gamma(t) = \int_0^t g\big(\tau, q_\gamma(\tau), \dot{q}_\gamma(\tau)\big)\, d\tau.$$

By assumption (A3) we suppose without loss of generality that the integrand $g(\cdot,\cdot,\cdot)$ is continuous on $T_\varepsilon \times H \times H$. Since

$$|\dot{q}(t) - \dot{q}_\gamma(t)| \le f\big(t, \sqrt{\varepsilon + \gamma}\big) + \varepsilon + \gamma, \quad t \in T_\varepsilon,$$

by the above estimates and since the function f can be assumed to be continuous on $T_\varepsilon \times I\!R_+$ by (A2), we get that the difference $|\vartheta(t) - \vartheta_\gamma(t)|$ is uniformly small on T provided that ε and γ are chosen to be sufficiently small. The latter conclusion and the estimate (4.21) imply that the *trajectory* $\big(q_\gamma(t), \vartheta_\gamma(t)\big)$ to the extended differential inclusion (4.4) is uniformly close to the *quasitrajectory* $w(t) = \big(q(t), \vartheta(t)\big)$ built is Step 1. By taking into account the result of Step 1, this completes the proof of the theorem. □

Next let us derive from the density result of Theorem 4.1 a *Bogolyubov-type theorem* for the MOSL differential inclusion (1.1) with the cost integral functional (4.1) under the assumptions (A1), (A2), and (A3). This theorem ensures *not only* the *uniform approximation* of relaxed trajectories to (1.1) by ordinary ones but also provides an important information on behavior of the *integral functional* $I[x]$ in (4.1) under such an approximation. To proceed, we consider the extended-real-valued function

$$g_F(t,x,v) := g(t,x,v) + \delta\big(v; F(t,x)\big),$$

where $\delta(\cdot;\Omega)$ stands for the *indicator function* of a set that is equal to 0 on the set and equal to ∞ outside the set. Define then

$$\widehat{g}(t,x,v) := (g_F)_v^{**}(t,x,v) \tag{4.22}$$

the *biconjugate/bypolar* function to $g_F(t,x,\cdot)$ with respect to *velocity*, i.e., the greatest, proper, *convex*, and lower semicontinuous function in v that is majorized by g_F.

Theorem 4.2. (Extended Bogolyubov Theorem for MOSL Differential Inclusions) *Under the assumptions of Theorem* 4.1 *the following hold:*
For every AC trajectory $\widetilde{x}(\cdot)$ *to the convexified differential inclusion* (2.4) *there is a sequence* $\{x_k(\cdot)\}$, $k \in I\!N$, *of AC trajectories to the original differential inclusion* (1.1) *such that*

$$\lim_{k\to\infty} \max_{t\in T} \big|x_k(t) - \widetilde{x}(t)\big| = 0, \tag{4.23}$$

$$\lim_{k\to\infty} \max_{t\in T} \Big| \int_0^t \Big[g\big(\tau, x_k(\tau), \dot{x}_k(\tau)\big) - \widehat{g}\big(\tau, \widetilde{x}(\tau), \dot{\widetilde{x}}(\tau)\big)\Big]\, d\tau \Big| = 0. \tag{4.24}$$

Proof. This can be derived from Theorem 4.1 similarly to the device in [4], where a (full) *Lipschitzian* analog of Theorem 4.1 was established and employed for compact-valued differential inclusions in separable Banach spaces.

Indeed, it is shown in [4] (the proof of this part holds with no change under our assumptions) that the pair $y(\cdot) = \big(x(\cdot), s(\cdot)\big)$ is a solution of the *convexified extended* differential inclusion (4.5) *if and only if* one has

$$\begin{cases} \dot{x}(t) \in \overline{\text{co}}\, F\big(t, x(t)\big) \;\text{ for a.e. }\; t \in T, \quad x(0) = x_0, \\ \dot{s}(t) = \widehat{g}\big(t, x(t), \dot{x}(t)\big) \;\text{ for a.e. }\; t \in T, \quad s(0) = 0. \end{cases} \tag{4.25}$$

Thus taking the designated solution $\widetilde{x}(\cdot)$ to the convexified differential inclusion (2.4) in the statement of the theorem, we consider the pair

$$\big(\widetilde{x}(t), \widetilde{s}(t)\big) \;\text{ with }\; \widetilde{s}(t) := \int_0^t \widehat{g}\big(\tau, \widetilde{x}(\tau), \dot{\widetilde{x}}(\tau)\big)\, d\tau, \quad t \in T,$$

which, by (4.25), is a solution to (4.5). Employing now Theorem 4.1, we find a sequence of solutions $y_k(\cdot) := \big(x_k(\cdot), s_k(\cdot)\big)$ to the extended inclusion (4.4) such that

$$x_k(t) \to \widetilde{x}(t) \;\text{ and }\; s_k(t) = \int_0^t g\big(\tau, x_k(\tau), \dot{x}_k(\tau)\big)\, d\tau \to \widetilde{s}(t) \;\text{ uniformly on }\; T \;\text{ as }\; k \to \infty.$$

The latter gives (4.23) and (4.24) and completes the proof of the theorem. □

Now, as a *bonus* of the technique developed in the proof of Theorem 4.1 combined with the proof of Theorem 3.2, we establish a version of Theorem 3.2 on the *strong* convergence of *discrete approximations* that does not require any additional (joint continuity) assumptions on $F(t, x)$ and uses only the standing assumptions (A1) and (A2). In particular, the following result allows us to deal with discrete approximations of *MOSL* differential inclusions and control systems whose initial data are merely *measurable* in time. This seems to be *new* (even for *fully Lipschitzian* problems with respect to state variables in finite-dimensional spaces) in the theory of discrete approximations and makes it possible to employ the *method of discrete approximations* as a *vehicle* for the qualitative and quantitative study of continuous-time systems with the *measurable dependence* on time variables, which was not the case in the previous developments and applications; see, e.g., [16, 17, 18] and the references therein.

Theorem 4.3. (Strong Convergence of Discrete Approximations for Nonconvex MOSL Differential Inclusions under Almost Continuity Assumptions) *Let the standing assumptions* (A1) *and* (A2) *be satisfied. Then for every AC solution $x(\cdot)$ to* (1.1) *there is a sequence of partitions Δ_k of T given in* (3.11) *and a sequence of piecewise linear solutions $z_k(\cdot)$ to the discretized inclusions* (3.2) *on Δ_k as $k \to \infty$ such that the strong convergence relationships* (3.12) *hold.*

Proof. Fix an AC solution $x\colon T \to H$ to the differential inclusion (1.1). Following the proof of (4.9) in Theorem 4.1, where the convex-valuedness of the mapping $\overline{\text{co}}\, G(\cdot,\cdot)$ does not play any role while its almost continuity is crucial, for any $\lambda > 0$ we find a compact subset $T_\lambda \subset T$ with $\text{mes}(T_\lambda) > 1 - \lambda^2$ and an absolutely continuous function $y\colon T \to H$ with the *piecewise constant derivative* such that the mappings $F(\cdot,\cdot)$ and $f(\cdot,\cdot)$ from (A1) and (A2) are *continuous* on $T_\lambda \times H$ and the estimates

$$\text{dist}\big(\dot{y}(t); F(t, y(t))\big) \le \lambda \ \text{ on } \ T_\lambda \ \text{ and } \ \|\dot{y} - \dot{x}\|_{L^1(T;H)} \le \lambda \tag{4.26}$$

are satisfied. Thus there is a subdivision

$$\Lambda_m := \big\{0 = \tau_0^m < \tau_1^m < \ldots < \tau_m^m = 1\big\}, \quad m \in I\!N,$$

of T such that $\dot{y}(t)$ is piecewise constant on every subinterval $[\tau_j^m, \tau_{j+1}^m)$, $j = 0, \ldots, m-1$. We can assume without loss of generality that $\tau_j^m \in T_\lambda$ for each $j \in \{0, \ldots, m-1\}$.

Note that the above functions $y(t) = y_\lambda(t)$ satisfying (4.26) are *not* feasible trajectories to the discretized inclusions (3.2). Now, arguing similarly to the proof of Theorem 3.2, we can *approximate* them *strongly* in $W^{1,p}(T;H)$, $p \in [1,\infty)$, by a sequence of piecewise linear *trajectories* $z_k(t)$ to the discrete inclusions (3.2) defined on the appropriate subintervals

$$\Delta_k := \big\{0 = t_0^k < t_1^k < \ldots < t_k^k = 1\big\} \ \text{ with } \ h_k := \max_{0 \le j \le k-1} \big\{t_{j+1}^k - t_j^k\big\} \downarrow 0 \ \text{ as } \ k \to \infty$$

of T. To proceed, we use the *uniform continuity* property of $F(\cdot,\cdot)$ on $T_\lambda \times H$ along the functions $y(t) = y_\lambda(t)$ from (4.26) meaning that for every $\varepsilon > 0$ there is $\eta > 0$ ensuring

$$d_H\big(F(t, y(t)), F(\tau, x)\big) \le \varepsilon \ \text{ whenever } \ t, \tau \in T_\lambda, \ |t - \tau| \le \eta, \ |y(t) - x| \le \eta,$$

and then employ the *projection method* as in the proof of Theorem 3.2, which is essentially based on the *MOSL property* of F. The reader can furnish all the details similarly to the proof of Theorem 3.2. □

Remark 4.4. (Differential Inclusions with Noncompact Values) Careful analysis and appropriate technical modifications of the given proofs for the above approximation and relaxation results show that the *compact-valuedness* requirement on $F(\cdot,\cdot)$ can be *dropped* under the basic assumptions (A1), (A2), and (A3). In particular, the *projection* constructions essentially used in the proofs above, which eventually require the compactness of underlying sets in infinite dimensions, can be replaced in the approximating procedures by *density results* of Lau's nearest point type; see [14].

5 Discrete Approximations of the Generalized Bolza Problem for MOSL Differential Inclusions

In this section we study discrete approximations of *dynamic optimization* problems over trajectories to MOSL differential inclusions. The main problem under consideration is known as the *generalized Bolza problem* and is described as follows:

$$\text{minimize } J[x] := \varphi\big(x(1)\big) + \int_0^1 g\big(t, x(t), \dot{x}(t)\big)\, dt \tag{5.1}$$

over AC trajectories $x\colon T \to H$ to the original differential inclusion (1.1) subject to the general *endpoint constraints*

$$x(1) \in \Omega \subset H. \tag{5.2}$$

This problem denoted as (P) has been well recognized as a basic model in dynamic optimization that covers both conventional and nonconventional problems of the (one-dimensional in time) *calculus of variations* and constrained *optimal control* for open-loop and closed-loop systems; see [2, 17, 20] for more discussions.

The cost functional $J[x]$ in (5.1) differs from $I[x]$ in (4.1) considered in Section 4 in connection with the extended Bogolyubov theorem by the endpoint (or *Mayer*) term $\varphi\big(x(1)\big)$ typical in problems of optimal control.

Our primary attention in this section is paid to constructing *well-posed discrete approximations* to problem (P) by a sequence of optimization problems governed by discrete inclusions whose optimal solutions *strongly* in $W^{1,p}(T;H)$, $p \in [1,\infty)$, converge to the *given optimal solution* $\bar{x}(t)$ for the continuous-time problem (P). More precisely (and more generally), we deal with the so-called "intermediate local minimizers" to (P) in the sense of [16], which are situated *strictly* between the classical weak and strong local minima; see [16] and [17, Subsection 6.1.2] for detailed discussions and examples.

Recall that a feasible trajectory $\bar{x}(\cdot)$ to (P) is an *intermediate local minimizer* (ILM) of rank $p \in [1,\infty)$ to this problem if there are numbers $\varepsilon > 0$ and $\alpha \ge 0$ such that $J[\bar{x}] \le J[x]$ for any feasible trajectory $x(\cdot)$ to (P) satisfying

$$\big|x(t) - \bar{x}(t)\big| < \varepsilon \;\text{ on }\; T \;\text{ and }\; \alpha \int_0^1 \big|\dot{x}(t) - \dot{\bar{x}}(t)\big|^p\, dt < \varepsilon. \tag{5.3}$$

The relationships in (5.3) actually mean that we consider a neighborhood of $\bar{x}(\cdot)$ in the Sobolev space $W^{1,p}(T;H)$. The case of $\alpha = 0$ in (5.3) corresponds to the classical *strong* local minimum and surely includes global solutions to (P) in the usual sense. The classical *weak* local minimum corresponds to (5.3) with $\alpha \neq 0$ and $p = \infty$, which is more restrictive.

In what follows we are going to construct strong discrete approximations of the local solution $\bar{x}(\cdot)$ in the afore-mentioned sense under *localizing* assumptions (A1), (A2), and (A3). This means that we need their fulfillment not on the whole space H as formulated but only on some bounded set $U \subset H$ with includes $\bar{x}(t)$, $\dot{\bar{x}}(t)$, and the underlying *neighborhood* of the intermediate local minimizer. Furthermore, for simplicity and convenience we slightly modify the assumptions in (A3) on the integral g in (5.1) requiring that

(A3′) $g(t,\cdot,\cdot)$ is continuous on $U\times U$ uniformly in $t\in T$, while $g(\cdot,x,v)$ is measurable on T and its absolute value is majorized by a summable function uniformly in $(x,v)\in U\times U$.

As well known, (A3′) implies the almost continuity property of $g(\cdot,\cdot,\cdot)$ in (A3) in *separable* spaces; so we can use the results obtained in Section 4 under (A3′) in separable Hilbert spaces. On the other hand, we can *avoid* the separability requirement on H if g is assumed to be *continuous* in t (i.e., jointly with respect to all its variables); see Remark 5.2. In fact, based on the technique developed in the proof of Theorem 4.1, one can proceed in the slightly modified construction below with the (localized) *almost continuity assumption* on g in *nonseparable* spaces as in (A3) including the integrand g into the discrete approximation procedure of Theorem 4.3; we leave details to the reader.

To proceed, we also need to add to (A1), (A2), and (A3′) the following unrestrictive assumptions concerning the new data φ and Ω in problem (P) and involving the aforementioned bounded set $U\subset H$:

(A4) The function $\varphi(\cdot)$ from (5.1) is continuous on U and the set Ω from (5.2) is closed around $\bar{x}(1)$.

Note that the results on discrete approximations obtained below significantly improved the known ones (see [16, 17, 18] with the discussions and references therein) in both finite-dimensional and Hilbert space settings by replacing the *full* Lipschitz continuity of $F(t,\cdot)$ by the weaker *MOSL property* and also by replacing of the strong continuity-like requirements with respect to t by the *almost continuity* assumptions on $F(\cdot,\cdot)$, which allows us to cover *measurable* in time data; see the above discussions. At the same time the *compactness* requirement on the set values $F(t,x)$ seems to be *essential* for the results of this section as well as for those in [16, 17, 18].

To proceed, we need some amount of *relaxation stability*. Similarly to [16, 17], let us formalize this requirement in the following way. Along with (P), consider the *relaxed* generalized Bolza problem (R) given by:

$$\text{minimize}\quad \widehat{J}[x]:=\varphi\big(x(1)\big)+\int_0^1 \widehat{g}\big(t,x(t),\dot{x}(t)\big)\,dt \tag{5.4}$$

subject to the *convexified* differential inclusion

$$\dot{x}(t)\in\overline{\text{co}}\,F\big(t,x(t)\big)\quad\text{for a.e.}\quad t\in T,\qquad x(0)=x_0 \tag{5.5}$$

with the endpoint constraints (5.2). We say that an absolutely continuous function $\bar{x}\colon T\to H$ is a *relaxed intermediate local minimizer* (RILM) of rank $p\in[1,\infty)$ to the original Bolza problem (P) if $\bar{x}(\cdot)$ is feasible to (P) and provides an intermediate local minimum of this rank to the relaxed problem (R) with the same cost value $J[\bar{x}]=\widehat{J}[\bar{x}]$.

Clearly that any RILM for (P) is ILM to this problem and that the opposite is true if this (P) is *convex* in the sense that the velocity sets $F(t,x)$ are convex and the integrand $g(t,x,v)$ is convex in the velocity variable v. Moreover, the latter property is satisfied *far beyond convexity*; see a number of sufficient conditions for it in [16, 17, 20] and the references therein. A new result in this direction follows from Theorem 4.2 and is used in what follows; see the proof of Theorem 5.1(iii).

Take and fix an arbitrary RILM $\bar{x}(\cdot)$ for the original problem (P) and suppose for convenience (and without loss of generality) that $p = 2$, that $\alpha = 1$, and that

$$\bar{x}(t) + \varepsilon/2 \in U \quad \text{whenever} \quad t \in T$$

for the constants (p, α, ε) in (5.3). We now construct in the following way a desired sequence of discretized problems (P_k) as $k \in I\!N$ whose *optimal solutions* exist and strongly approximate the given RILM $\bar{x}(\cdot)$ as $k \to \infty$.

Using Theorem 4.3, find a sequence of discrete partitions $\Delta_k = \{t_j \mid j = 0, \ldots, k\}$ of T as in (3.11)—omitting the upper index "k" for simplicity—and a sequence of piecewise linear solutions $\widetilde{z}_k(\cdot)$ to the discretized inclusions (3.2) such that the convergence relationships (3.12) hold with $x(\cdot) = \bar{x}(\cdot)$ and $z_k(\cdot) = \widetilde{z}_k(\cdot)$. Then problem (P_k) for each $k \in I\!N$ consists of minimizing the *cost functional*

$$\begin{aligned} J_k[z] := \quad & \varphi\big(z(t_k)\big) + \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} g\Big(t, z(t_j), \frac{z(t_{j+1}) - z(t_j)}{t_{j+1} - t_j}\Big)\, dt \\ & + \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \Big| \frac{z(t_{j+1}) - z(t_j)}{t_{j+1} - t_j} - \dot{\bar{x}}(t) \Big|^2 dt \end{aligned} \tag{5.6}$$

over piecewise linear trajectories $z(\cdot)$ to the discretized inclusion (3.2) subject to the *state* and *endpoint constraints*

$$\big|z(t_j) - \bar{x}(t_j)\big| \le \frac{\varepsilon}{2} \quad \text{for all} \quad j = 1, \ldots, k, \tag{5.7}$$

$$\sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \Big| \frac{z(t_{j+1}) - z(t_j)}{t_{j+1} - t_j} - \dot{\bar{x}}(t) \Big|^2 dt \le \frac{\varepsilon}{2}, \tag{5.8}$$

$$z(t_k) \in \Omega + \eta_k I\!B \quad \text{with} \quad \eta_k := \big|\widetilde{z}_k(t_k) - \bar{x}(t_k)\big|, \tag{5.9}$$

where $\eta_k \downarrow 0$ as $k \to \infty$ by Theorem 4.3 employed for $x(\cdot) = \bar{x}(\cdot)$ and $z_k(\cdot) = \widetilde{z}_k(\cdot)$.

The following major result ensures the *strong* $W^{1,p}$-approximation of *any given RILM* $\bar{x}(\cdot)$ to (P) by optimal solutions to the discrete problems (P_k) and, furthermore, justifies such a discrete approximation for an *arbitrary strong local minimizer* to the original Bolza (P) with no endpoint constraints (5.2).

Theorem 5.1. (Strong Convergence of Discrete Optimal Solutions to RILMs and Strong Local Minimizers for the Bolza Problem). *Let $\bar{x}(\cdot)$ be a RILM to the Bolza problem (P) under the localized assumptions* (A1), (A2), (A3′), *and* (A4) *in separable Hilbert spaces H. The following assertions hold:*

(i) *Each discrete approximation problem (P_k) admits an optimal solution.*

(ii) *Any sequence of optimal solutions $\{\bar{z}_k(\cdot)\}$ to (P_k) converges to $\bar{x}(\cdot)$ strongly in the space $W^{1,p}(T; H)$ as $p \in [1, \infty)$.*

(iii) *If $\Omega = H$ in (P), then the above conclusions of the theorem are fulfilled for an arbitrary strong local minimizer $\bar{x}(\cdot)$ to the original problem.*

Proof. To justify (i), we first observe that the set of *feasible* solutions to each problem (P_k) is *nonempty* for all $k \in I\!N$ sufficiently large. Indeed, that approximating trajectories $\widetilde{z}_k(\cdot)$ are feasible to (P_k) as $k \to \infty$ due to Theorem 4.3 and the construction of (P_k). This observation holds for *any ILM* $\bar{x}(\cdot)$ by its definition in (5.3). Then the *existence of optimal solutions* to (P_k) in assertions (i) and (iii) follows directly from the classical Weierstrass existence theorem due to the compactness and continuity (in x) assumptions imposed on the initial data of (P).

Next we prove (ii). It is easy to see (from the proof of Theorem 5.1) that without loss of generality the knots t_j in (P_k) can be chosen as points of continuity of the velocity mapping $F(t,\cdot)$. Let us first check that

$$J_k[\widetilde{z}_k] \to J[\bar{x}] \quad \text{as} \quad k \to \infty \tag{5.10}$$

along some subsequence of $k \in I\!N$ for the cost functionals (5.1) and (5.6) in problems (P_k) and (P), respectively, where $\bar{x}(\cdot)$ and $\widetilde{z}_k(\cdot)$ are related by Theorem 4.3. Since φ is continuous around $\bar{x}(1)$, the convergence relation (5.10) obviously reduces to

$$\sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} g\Big(t, \widetilde{z}_k(t_j), \frac{\widetilde{z}_k(t_{j+1}) - \widetilde{z}_k(t_j)}{t_{j+1} - t_j}\Big)\, dt \to \int_0^1 g\big(t, \bar{x}(t), \dot{\bar{x}}(t)\big)\, dt \quad \text{as} \quad k \to \infty,$$

which follows from Theorem 4.3 and Lebesgue's dominated convergence theorem for the Bochner integral that is valid under (A3′).

The arguments above did not involve the property of $\bar{x}(\cdot)$ to be a *relaxed* ILM to the Bolza problem (P). Now, employing this property and taking *any* sequence $\{\bar{z}_k(\cdot)\}$ of optimal solutions to the discrete problems (P_k), let us show that

$$\lim_{k \to \infty} \int_0^1 \big|\dot{\bar{z}}_k(t) - \dot{\bar{x}}(t)\big|^2\, dt = 0, \tag{5.11}$$

which obviously implies the conclusion in (ii). Assuming the contrary and using the Dunford theorem on the *weak precompactness* in $L^1(T;H)$ (see, e.g., [6, Theorem IV.1]), we find $\gamma > 0$ and $v(\cdot) \in L^1(T;H)$ such that

$$\int_0^1 \big|\dot{\bar{z}}_k(t) - \dot{\bar{x}}(t)\big|^2\, dt \to \gamma \quad \text{and} \quad \dot{\bar{z}}_k(\cdot) \to v(\cdot) \quad \text{weakly in} \quad L^1(T;H) \tag{5.12}$$

along a subsequence of $k \in I\!N$, which we identify as usual with the whole natural series. Since the Bochner integral is a linear continuous operator from $L^1(T;H)$ into H, it remains continuous with respect to the weak topology. Taking also into account Lemma 2.2(ii) on the precompactness in $\mathcal{C}(T;H)$ of the solution set to (3.2) under the assumptions made, we find an absolutely continuous function $\widetilde{x}\colon T \to H$ such that

$$\widetilde{x}(t) = x_0 + \int_0^1 v(\tau)\, d\tau \quad \text{for all} \quad t \in T,$$

and thus $\dot{\widetilde{x}}(t) = v(t)$ for a.e. $t \in T$ and $\dot{\bar{z}}_k(\cdot) \to \dot{\widetilde{x}}(\cdot)$ weakly in $L^1(T;H)$ by (5.12) as $k \to \infty$.

Observe furthermore that the limiting function $\widetilde{x}(\cdot)$ is a solution to the *convexified* differential inclusion (5.5). Indeed, it follows from the classical Mazur theorem that weak convergence of $\{\dot{\bar{z}}_k(\cdot)\}$ from (5.12) implies the strong in $L^1(T;H)$ convergence to $\dot{\widetilde{x}}(\cdot)$ of some *convex combinations* of $\dot{\bar{z}}_k(\cdot)$, and hence the *a.e. pointwise* convergence to $\dot{\widetilde{x}}(t)$ of (a subsequence of) these convex combinations. Thus inclusion (5.5) for $\widetilde{x}(\cdot)$ follows from those in (3.2) for all $\bar{z}_k(\cdot)$ as $k \to \infty$. By passing to the limit in the constraint relationships (5.7) and (5.9) for $\bar{z}_k(\cdot)$, we conclude that

$$|\widetilde{x}(t) - \bar{x}(t)| \le \varepsilon/2 \;\text{ on }\; T \;\text{ and }\; \widetilde{x}(1) \in \Omega.$$

For passing to the limit in (5.8), observe that the integral functional

$$I[u] := \int_0^1 |u(t) - \dot{\bar{x}}(t)|^2 \, dt$$

is *lower semicontinuous* in the *weak* topology of $L^2(T;H)$ due to the *convexity* of the integrand in u. Since the weak convergence of $\dot{\bar{z}}_k(\cdot) \to \dot{\widetilde{x}}(\cdot)$ in $L^1(T;H)$ is equivalent to the one in $L^2(T;H)$ by the unform boundedness property of Lemma 2.1, we conclude from the afore-mentioned lower semicontinuity and the piecewise linear structure of $\bar{z}_k(\cdot)$ that the limiting function $\widetilde{x}(\cdot)$ satisfies the integral constraint in (5.3), and thus it belongs to prescribed ε-neighborhood of the RILM $\bar{x}(\cdot)$ under consideration.

Since the approximating trajectories $\widetilde{z}_k(\cdot)$ from Theorem 4.3 are *feasible* to (P_k) while $\bar{z}_k(\cdot)$ are *optimal* to these problems as $k \to \infty$, we have

$$J_k[\bar{z}_k] \le J_k[\widetilde{z}_k] \;\text{ for all large }\; k \in I\!N. \tag{5.13}$$

Taking into account the structure of J_k in (5.6) and the arguments above, as well as construction (4.22) of the convexified integrand $\widehat{g}$ in (5.4), we get from (5.10) by passing to the limit in (5.13) that

$$\varphi(\widetilde{x}(1)) + \int_0^1 \widehat{g}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)) \, dt + \gamma \le J[\bar{x}],$$

where $\gamma > 0$ by (5.12). Thus we arrive at the contradiction

$$\widehat{J}[\widetilde{x}] < J[\bar{x}] = \widehat{J}[\bar{x}]$$

to the fact that $\bar{x}(\cdot)$ is a RILM to (P), which therefore justifies (5.11) and completes the proof of assertion (ii) in the theorem.

It remains to prove the convergence statement in (iii) for an arbitrary *strong local minimizer* $\bar{x}(\cdot)$ to the original Bolza problem (P) with no endpoint constraints (5.2). It turns out that in this case, under the assumptions of the theorem for MOSL differential inclusions, any *strong* local minimizer to (P) is a strong local minimizer for the *relaxed* problem (R), and hence it is a RILM to (P) enjoying the conclusion in (ii). Indeed, given a strong local minimizer $\bar{x}(\cdot)$ to (P) and assuming the contrary, for any $\varepsilon > 0$ we find a trajectory $\widetilde{x}(\cdot)$ to the convexified inclusion (5.5) such that

$$|\widetilde{x}(t) - \bar{x}(t)| < \varepsilon \;\text{ whenever }\; t \in T \;\text{ and}$$

$$\widehat{J}[\widetilde{x}] < \widehat{J}[\bar{x}] \le J[\bar{x}],$$

where the latter inequality is automatic. Now applying the relaxation result from Theorem 4.2 to the designated relaxed trajectory $\widetilde{x}(\cdot)$ and taking into account the continuity assumption on the cost function φ, we find a sequence of AC trajectories $x_k(\cdot)$ to the original inclusion (1.1) such that $x_k(t) \to \widetilde{x}(t)$ as $k \to \infty$ uniformly on T and

$$\liminf_{k\to\infty} J[x_k] \le \widehat{J}[\widetilde{x}] < J[\bar{x}]. \tag{5.14}$$

Note that all $x_k(\cdot)$ are feasible to (P)—by the absence of endpoint constraints—and belong to any prescribed neighborhood of $\bar{x}(\cdot)$ in the space $\mathcal{C}(T;H)$ for all $k \in I\!N$ sufficiently large. Thus (5.14) clearly contradicts the strong local minimality of $\bar{x}(\cdot)$ to the original problem (P). This completes the proof of assertion (iii) and of the whole theorem. □

Remark 5.2. (Simplified Discrete Approximations of the Bolza Problem with Continuous Integrands) Note that if the integrand g in (5.1) is assumed to be *continuous* in t, then the second term in representation (5.6) of the discretized cost functions $J_k[z]$ can be *simplified* in the constructions and conclusions of Theorem 5.1 by

$$\sum_{j=0}^{k-1}\Big(\frac{1}{t_{j+1}-t_j}\Big)g\Big(t_j, z(t_j), \frac{z(t_{j+1})-z(t_j)}{t_{j+1}-t_j}\Big) \tag{5.15}$$

for *any* discrete partition Δ_k of T from (3.11). Moreover, in this case we do *not* need to assume that the space H is *separable* in Theorem 5.1. This observation follows directly from the proof of Theorem 5.1 by using Theorem 3.2 instead of Theorem 4.3 therein.

Finally in this section, we obtain a general theorem on the *value convergence* of discrete approximations for *MOSL* differential inclusions extending previous results in this direction known for *full* Lipschitzian counterparts; see [17] and the references therein.

Observe that the cost functional (5.6) as well as constraints (5.7)–(5.9) in the discrete approximation problems (P_k) *explicitly* contain the *given* local minimizer $\bar{x}(\cdot)$ to the original problem (P). From the numerical viewpoint, it is important to construct discrete approximations involving only *initial data* of (P) but not information about its (local) optimal solutions, which may not even exist. To proceed in this way, we modify (P_k) considering instead it the the following sequence of discrete approximation problems $(\widetilde{P}_k)$:

$$\text{minimize } \ \widetilde{J}_k[z] := \varphi\big(z(1)\big) + \sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}} g\Big(t, z(t_j), \frac{z(t_{j+1})-z(t_j)}{t_{j+1}-t_j}\Big)\,dt$$

subject to the discretized inclusions (3.2) with the *perturbed* endpoint constraints (5.9), where the sequence η_k is not yet specified. Similarly to (5.15), we can simplify the approximating functional $\widetilde{J}_k$ if the integrand g is continuous in t. Denote

$$\inf(P), \quad \inf(R), \quad \text{and} \quad \widetilde{J}^0_k := \inf(\widetilde{P}_k) \quad \text{as} \quad k \in I\!N$$

the *optimal values* of the cost functionals in the original, relaxed, and discretized problems under consideration. We say that problem (P) is *stable with respect to relaxation* if

$$\inf(P) = \inf(R). \tag{5.16}$$

The reader can find a number of efficient conditions ensuring this property in [2, 4, 13, 16, 17] and the references therein.

The following theorem shows that the relaxation stability (5.16) is *necessary and sufficient* for the value convergence of discrete approximations for MOSL differential inclusions under appropriate perturbations of the endpoint constraints.

Theorem 5.3. (Value Convergence of Discrete Approximations for MOSL Differential Inclusions) *Let U be an open and bounded subset of a separable space H such that $x_m(\cdot) \in U$ as $t \in T$ and $m \in I\!N$ for a minimizing sequence of feasible solutions to (P). Suppose that the localized assumptions* (A1), (A2), (A3$'$), *and* (A4) *are satisfied whenever $(x, v) \in U \times U$ with Ω to be fully closed in* (A4). *Then the following assertions hold:*

(i) *There is a sequence of the endpoint constraint perturbations $\eta_k \downarrow 0$ in* (5.9) *such that*

$$\inf(R) \le \liminf_{k\to\infty} \widetilde{J}_k^0 \le \limsup_{k\to\infty} \widetilde{J}_k^0 \le \inf(P), \tag{5.17}$$

and so the relaxation stability (5.16) *ensures the value convergence* $\inf(\widetilde{P}_k) \to \inf(P)$ *of the above discrete approximations.*

(ii) *Conversely, the relaxation stability of (P) is also necessary for the value convergence* $\inf(\widetilde{P}_k) \to \inf(P)$ *of the discrete approximations with arbitrary perturbations $\eta_k \downarrow 0$ of the endpoint constraints.*

Proof. To justify (i), we take the minimizing sequence of feasible trajectories $x_m(\cdot)$ to (P) specified in the theorem and apply to each $x_m(\cdot)$ Theorem 4.3 on the strong approximation by discrete trajectories. Employing the standard *diagonal process*, we construct the trajectories $\widetilde{z}_k(\cdot)$ to the discretized inclusions (3.2) such that

$$\eta_k := \big|\widetilde{z}_k(1) - x_{m_k}(1)\big| \to 0 \ \text{ as } \ k \to \infty. \tag{5.18}$$

Then the proof of (5.17) is similar to the ones in assertions (i) and (ii) of Theorem 5.1 with the endpoint perturbations η_k specified in (5.18).

To justify the *converse* assertion (ii) in the theorem, we first observe that the relaxed problem (R) admits an *optimal solution* under the assumptions made. This follows from the compactness assertion (i) of Lemma 2.2 and the lower semicontinuity arguments in the proof of assertion (ii) of Theorem 5.1. Taking an optimal solution $\bar{x}(\cdot)$ to problem (R), we approximate it by feasible trajectories $\widetilde{x}_m(\cdot)$, $m \in I\!N$, to the original problem (P) in the sense of Theorem 4.2 and then strongly in $W^{1,2}(T; H)$ approximate each $\widetilde{x}_m(\cdot)$ by some trajectories $\widehat{z}_{m_k}(\cdot)$, $k \in I\!N$, to the discretized inclusions (3.2). Using again the diagonal process, we thus build the corresponding trajectories $\widehat{z}_k(\cdot)$ to (3.2) approximating $\bar{x}(\cdot)$ in the sense of Theorem 4.2 and define the endpoint perturbations η_k by

$$\eta_k := \big|\widehat{z}_k(1) - \bar{x}(1)\big| \to 0 \ \text{ as } \ k \to \infty. \tag{5.19}$$

Suppose now that (P) is *not stable with respect to relaxation*, i.e.,

$$\widehat{J}[\bar{x}] = \min(R) < \inf(P). \tag{5.20}$$

for the fixed optimal solution $\bar{x}(\cdot)$ to (R). Then we construct the discrete approximation problems $(\widetilde{P}_k)$ as above with the endpoint perturbations η_k specified in (5.19). By (5.19), the afore-mentioned approximating trajectories $\widehat{z}_k(\cdot)$ are *feasible* to (P_k). It follows from the construction of these trajectories and the assumed strict inequality in (5.20) that

$$\begin{aligned}\liminf_{k\to\infty} \widetilde{J}_k^0 &\le \liminf_{k\to\infty}\Big[\varphi\big(\widehat{z}_k(1)\big) + \int_0^1 g\big(t, \widehat{z}_k(t), \dot{\widehat{z}}(t)\big)\,dt\Big]\\ &\le \varphi\big(\bar{x}(1)\big) + \int_0^1 \widehat{g}\big(t, \bar{x}(t), \dot{\bar{x}}(t)\big)\,dt < \inf(P),\end{aligned}$$

which shows that the value convergence $\inf(\widetilde{P}_k) \to \inf(P)$ does not hold for the constructed sequence of discrete approximations. This completes the proof of theorem. $\square$

As in Remark 5.2, observe that Theorem 5.3 holds in *nonseparable* spaces H and the discrete approximation in $(\widetilde{P_k})$ can be *simplified* by (5.15) if the integrand g is assumed to be *continuous in time*. This follows from the application of Theorem 3.2 in the proof above.

Remark 5.4. (Value Convergence and Strong Solution Convergence of Semi-Discrete Approximations for MOSL Differential Inclusions). Similarly to the proofs of Theorem 5.1 and Theorem 5.3, we can establish the strong solution convergence and value convergence results for *semi-discrete approximations* of the generalized Bolza problem (P) under the same assumptions. To justify this, it is sufficient to proceed as in the proofs of the corresponding discrete approximation theorems with replacing there the application of Theorem 4.3 and Theorem 3.2 by that of Theorem 3.1 with *no additional separability or time-continuity assumptions.*

References

[1] Z. Artstein, First-order approximations for differential inclusions, Set-Valued Analysis 2 (1994) 7–17.

[2] J.-P. Aubin, A. Cellina, Differential Inclusions, Grundlehren Series (Fundamental Principles of Mathematical Sciences), vol. 264, Springer-Verlag, Berlin, 1984.

[3] N.N. Bogolyubov, Sur quelques methodes nouvelles dans le calculus des variations, Ann. Math. Pura Appl. 7 (1930) 249–271.

[4] F. De Blasi, G. Pianigiani, A.A. Tolstonogov, A Bogolyubov-type theorem with a nonconvex constraint in a Banach space, SIAM J. Control Optim. 43 (2004) 466–476.

[5] K. Deimling, Multivalued Differential Equations, De Grujter, Berlin, 1992.

[6] J. Diestel, J. J. Uhl, Jr., Vector Measures, American Mathematical Society, Providence, Rhode Island, 1977.

[7] T. Donchev, Functional differential inclusion with monotone right-hand side, Nonlinear Analysis 16 (1991) 533–542.

[8] T. Donchev, Semicontinuous differential inclusions, Rend. Sem. Mat. Univ. Padova 101 (1999) 147–160.

[9] T. Donchev, Approximation of lower semicontinuous differential inclusions, Numerical Funct. Analysis Optim. 22 (2001) 55–67.

[10] T. Donchev, E. Farkhi, Stability and Euler approximations of one-sided Lipschitz convex differential inclusions, SIAM J. Control Optim. 36 (1998) 780–796.

[11] T. Donchev, E. Farkhi, Euler approximation of discontinuous one-sided Lipschitz convex differential inclusions, in: A. Ioffe et al. (eds.), Calculus of Variations and Differential Equations, Chapman and Hall/CRC, New York, 1999, pp. 101–118.

[12] A.L. Dontchev, F. Lempio, Difference methods for differential inclusions: a survey, SIAM Reviews 34 (1992) 263–294.

[13] A.D. Ioffe, V.M. Tikhomirov, Theory of Extremal Problems, North-Holland, Amsterdam, 1979.

[14] K. Lau, Almost Chebyshev subsets in reflexive Banach spaces, Indiana Univ. Math. J. 27 (1978) 791–795.

[15] F. Lempio, V. Veliov, Discrete approximations of differential inclusions, Bayreuter Mathematische Schiften 54 (1998) 149–232.

[16] B.S. Mordukhovich, Discrete approximations and refined Euler-Lagrange conditions for differential inclusions, SIAM J. Control Optim. 33 (1995) 882–915.

[17] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation, II: Applications, Grundlehren Series (Fundamental Principles of Mathematical Sciences), vol. 331, Springer-Verlag, Berlin, 2006.

[18] B.S. Mordukhovich, Variational analysis of evolution inclusions, SIAM J. Optim., to appear.

[19] A. Plis, Trajectories and quasitrajectories of an orientor field, Bull. Acad. Polol. Sci., Ser. Sci. Math. Astron. Phys. 11 (1963) 369–370.

[20] A.A. Tolstonogov, Differential Inclusions in a Banach Space, Series on Mathematics and Its Applications, vol. 524, Kluwer, Dordrecht, 2000.